



Approximations de la projection perspective : modèle combiné et étude des singularités homographiques

Diane Lingrand

► To cite this version:

Diane Lingrand. Approximations de la projection perspective : modèle combiné et étude des singularités homographiques. RR-3682, INRIA. 1999. inria-00072989

HAL Id: inria-00072989

<https://inria.hal.science/inria-00072989>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Approximations de la projection perspective :
modèle combiné et étude des singularités
homographiques.***

Diane Lingrand

N° 3682

Avril 1999

_____ THÈME 3 _____



***apport
de recherche***



Approximations de la projection perspective : modèle combiné et étude des singularités homographiques.

Diane Lingrand

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet RobotVis

Rapport de recherche n° 3682 — Avril 1999 — 33 pages

Résumé : Les approximations de la projection perspective étudiées dans ce rapport, projections para-perspective, ortho-perspective et orthographique, peuvent être considérées à la fois comme des approximations géométriques et des approximations numériques (développement limité d'ordre précisé). Nous avons examiné les simplifications apportées par ces projections, notamment pour l'étude de la disparité vectorielle.

Ces modèles s'appliquent dans le cadre de la focalisation d'attention dans la mesure où ils rendent compte d'une zone locale ou fovéale de l'image observée et permettent une analyse numériquement robuste de la structure et du mouvement de la scène même avec un faible nombre de données.

Après avoir dérivé la matrice fondamentale dans tous ces cas, on s'est plus particulièrement intéressé aux cas de mouvement et de structure de la scène pour lesquels on obtient une relation homographique entre les points images et on a mis en évidence les propriétés géométriques et de mouvement correspondant aux cas de la projection para-perspective et orthographique.

Mots-clés : projection perspective, projection para-perspective, projection orthographique, projection ortho-perspective, homographie

Perspective projection's approximations : combined model and homographic singularities

Abstract: Perspective projection approximation studied in this report, para-perspective projection, ortho-perspective, and orthographic may be considered both as geometric and numerical (limited order expansions) specifications, as demonstrated here, analyzing which kind of simplifications are obtained using these projections, considering motion cues, in particular retinal disparity.

These restrictive models are especially useful when considering focus of attention since they correctly represent visual data in a small area of the retina, e.g. the fovea, while they yield robust numerical analysis of structure and motion of the scene even using a small amount of data.

After a generalization of the fundamental matrix to these projection models we have analyzed in which conditions image points could be related by a collineation, both as far as geometry and kinematic is concerned, when para-perspective or orthographic projection is considered.

Key-words: perspective projection, para-perspective, weak perspective, ortho-perspective, homography

1 Introduction

Dans le cadre de la modélisation des caméras à base de CCD utilisées en Vision par Ordinateur, la projection perspective est le modèle le plus couramment utilisé, car le plus expérimentalement valide, si on néglige les phénomènes non linéaires (distorsions, astigmatisme, ...) des caméras. Néanmoins, des approximations de ce modèle sont parfois tout aussi valides et donc à utiliser dans certains cas car elles fournissent alors des expressions bien plus simples donc soit solubles analytiquement, soit au moins plus robustes numériquement.

Par ailleurs, dans le cadre d'une étude générale sur les singularités des équations du problème de l'analyse du mouvement, nous avons été amenés à distinguer les cas où on peut utiliser la matrice fondamentale de ceux pour lesquels il faut avoir recours à une homographie[Fau93]. En effet, dans le cas de la projection perspective, il est connu depuis longtemps que la matrice fondamentale ne peut être utilisée lorsque le mouvement est une rotation pure ou lorsqu'il s'agit du mouvement d'un plan. En examinant les approximations de la projection perspective, on s'est aperçu que cette propriété n'était pas exactement vérifiée. L'objet de ce rapport est donc de déterminer des propriétés équivalentes.

Nous nous plaçons dans le cas de deux images acquises par des caméras non calibrées et considérons que des points d'intérêts dans les deux images ont été détectés et mis en correspondance les uns avec les autres (par exemple selon la méthode de [ZDFL94]).

Notations

les vecteurs et matrices sont notés en **gras**

les vecteurs sont représentés par des minuscules

les matrices sont représentées par des majuscules

$\mathbf{x} \wedge \mathbf{y}$ représente le produit vectoriel des vecteurs \mathbf{x} et \mathbf{y}

$\tilde{\mathbf{x}}$ représente la matrice antisymétrique de l'opérateur produit vectoriel du vecteur \mathbf{x}

$\tilde{\mathbf{x}} \mathbf{y}$ représente alors le produit vectoriel des vecteurs \mathbf{x} et \mathbf{y}

$\mathbf{x} \cdot \mathbf{y}$ représente le produit scalaire des vecteurs \mathbf{x} et \mathbf{y}

$\mathbf{x}^T \mathbf{y}$ représente le produit scalaire des vecteurs \mathbf{x} et \mathbf{y}

2 La projection perspective

Le modèle sténopé, appelé également "projectif linéaire", consiste à considérer que la caméra effectue une projection perspective de centre C sur un plan rétinien Π . Afin d'exprimer cette projection dans les coordonnées de l'espace et de l'image, il est nécessaire de connaître le changement de repère entre la scène observée et le repère attaché à la caméra. Un point \mathbf{M} de l'espace a pour coordonnées (X, Y, Z) dans le repère \mathcal{R} de l'espace. Un point \mathbf{m} de l'image a pour coordonnées (u, v) dans le repère image \mathcal{R}_i , déterminé par les axes horizontal et vertical de l'image avec pour origine le coin supérieur gauche. Ce repère image induit un repère espace \mathcal{R}' dont le troisième axe est donné par l'axe optique et le centre

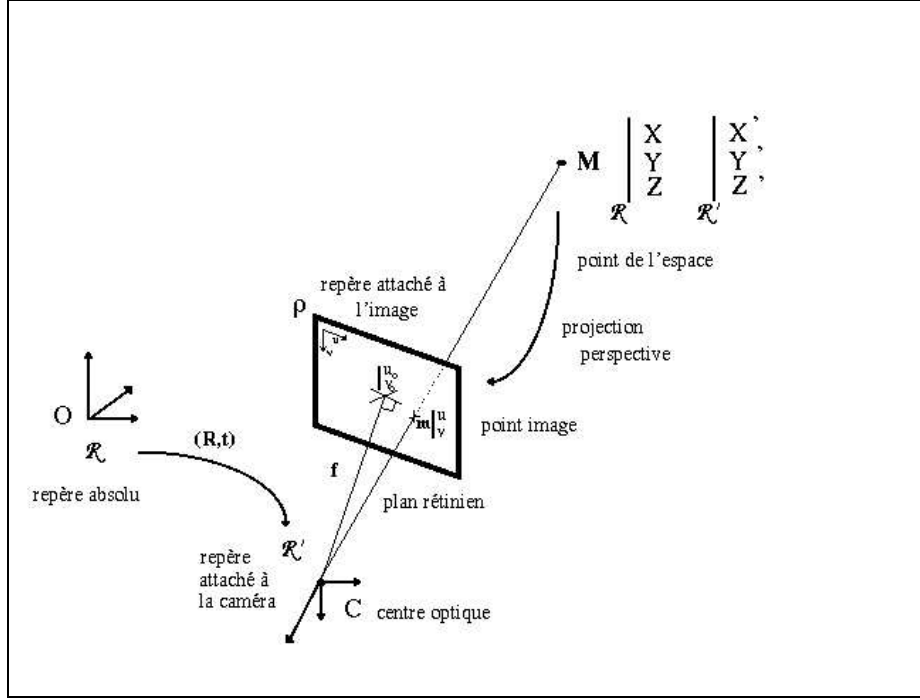


FIG. 1 – Le modèle sténopé

par le centre optique C tandis que les coordonnées X et Y sont alignées avec les lignes et colonnes de l'image. Dans le repère image \mathcal{R}_i , le point C a pour coordonnées $(u_0, v_0, -f)$, où f est la distance focale (distance du plan rétinien au centre optique). Le point M de l'espace a pour coordonnées (X', Y', Z') dans le repère \mathcal{R}' . La transformation entre les repères \mathcal{R} et \mathcal{R}' est la composée d'une rotation \mathbf{R} et d'une translation \mathbf{t} . La figure 1 clarifie les notations.

La projection perspective (voir figure 2) est représentée par l'équation suivante, dans le repère attaché à la caméra :

$$Z \mathbf{m} = \begin{pmatrix} \alpha_u & \gamma & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{M} \quad (1)$$

Dans cette projection,

α_u et α_v représentent les tailles horizontale et verticale des pixels. On décompose parfois $\alpha_u = f k_u$ et $\alpha_v = f k_v$ afin d'exhiber que ces variables contiennent la focale. Lorsque α_u et α_v diffèrent, les pixels ne sont pas carrés mais rectangulaires.

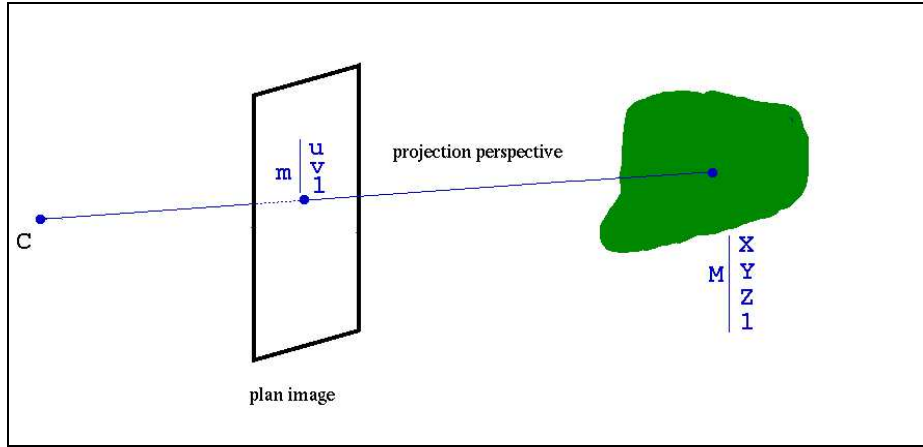


FIG. 2 – Principe de la projection projective

u_0 et v_0 représentent la projection du centre optique sur la rétine, le centre de la zone fovéale.

γ mesure le défaut d'orthogonalité du plan des cellules CCD à par rapport à l'axe optique. Si γ n'est pas nul, les pixels, au lieu d'être rectangulaires, ont une forme de parallélogrammes.

3 Les approximations de la projection perspective

Les approximations de la projection perspective sont au nombre de trois :

- projection para-perspective
- projection ortho-perspective
- projection affine, appelée aussi orthographique, perspective faible (*“weak perspective”*) ou orthographique graduée (*“scaled orthographique”*) selon son paramétrage.

Ces types de projection diffèrent par le nombre de paramètres qu'ils comportent et par leur plus ou moins bonne adéquation aux données observées. Leur analyse est l'objet des paragraphes suivants.

3.1 État de l'art

Bien que, généralement, le modèle de projection perspective soit le plus souvent utilisé dans la littérature, différentes approximations ont été utilisées notamment par Aloimonos [Alo90], Dementhon et Davis [DD89], Horaud et Christy [HCD94, HDLC97], Soatto et Perona [SP95] ainsi que Lingrand et Viéville [LV96].

Poelman et Kanade [PK94] ont utilisé le modèle para-perspectif sur des séquences d'images et trouvent des résultats meilleurs que pour le modèle orthographique lorsque le mouvement de translation suivant l'axe optique ou orthogonal à l'axe optique est important. Cependant, le modèle para-perspectif nécessite de connaître le centre optique ainsi que la focale. Si ces paramètres ne sont pas connus, alors, le modèle orthographique est équivalent.

Boufama et Weinshall [BWW94] ont effectué une comparaison des modèles de projection orthographique et perspectif avec des algorithmes utilisant des invariants sur trois séquences d'images avec différentes tailles de champ de vue. Le modèle orthographique s'avérant plus performant pour un petit champ de vue et le modèle perspectif plus performant pour un large champ de vue, il vient naturellement à l'esprit des auteurs d'utiliser un algorithme combinant ces deux approches. Ce type d'approche par stratification a été utilisé par d'autres auteurs pour le mouvement par exemple [HCM97].

Lingrand et Viéville, [LV95], ont utilisé le modèle orthographique dans le cas où le mouvement de la caméra conserve son plan rétinien et d'une vision fovéale pour permettre de générer des équations de reconstruction plus robustes numériquement. Quan [Qua96] a utilisé ce même modèle mais pour des séquences d'images.

Dans la suite, on se placera dans le repère \mathcal{R}' attaché à la caméra et on notera $\mathbf{M} = [X, Y, Z, \{0, 1\}]$ un point de l'espace dans ce repère respectivement à l'infini et à distance finie.

3.2 La projection para-perspective

La projection para-perspective a été utilisée pour la première fois par Ohta [OMS81] afin de résoudre un problème de reconstruction 3D à partir de textures (*shape from texture*). Elle fut ensuite réutilisée par divers auteurs, notamment [CH94].

Ce modèle est bien adapté à la focalisation d'attention autour d'un angle de vue particulier.

3.2.1 Approximation géométrique

Cette approximation peut être vue comme une approximation géométrique, c'est-à-dire en supposant que l'objet est loin de la caméra ou proche de l'axe optique. La projection est alors correctement approximée par la composée d'une projection parallèle sur un plan orthogonal à l'axe optique et d'une projection perspective. La direction de la projection parallèle dépend de la scène observée (ou de l'objet observé) et est donnée par la direction du rayon partant du centre optique de la caméra vers le centroïde de la scène, i.e. le point de profondeur moyenne. Le plan image fictif sur lequel sont projetés les points issus de la projection parallèle est parallèle au plan rétinien et passe par le centroïde. La figure 3 illustre cette situation qui consiste à regarder "en coin".

Notons $\mathbf{M}_0 = [X_0 \ Y_0 \ Z_0 \ 1]^T$ le point de coordonnées moyennes (et donc de profondeur moyenne). Le plan image auxiliaire a pour équation $Z = Z_0$. Pour un point $\mathbf{M} = [X \ Y \ Z \ 1]^T$ de l'espace, notons $\mathbf{M}_a = [X_a \ Y_a \ Z_0 \ 1]^T$ le point image auxiliaire correspondant.

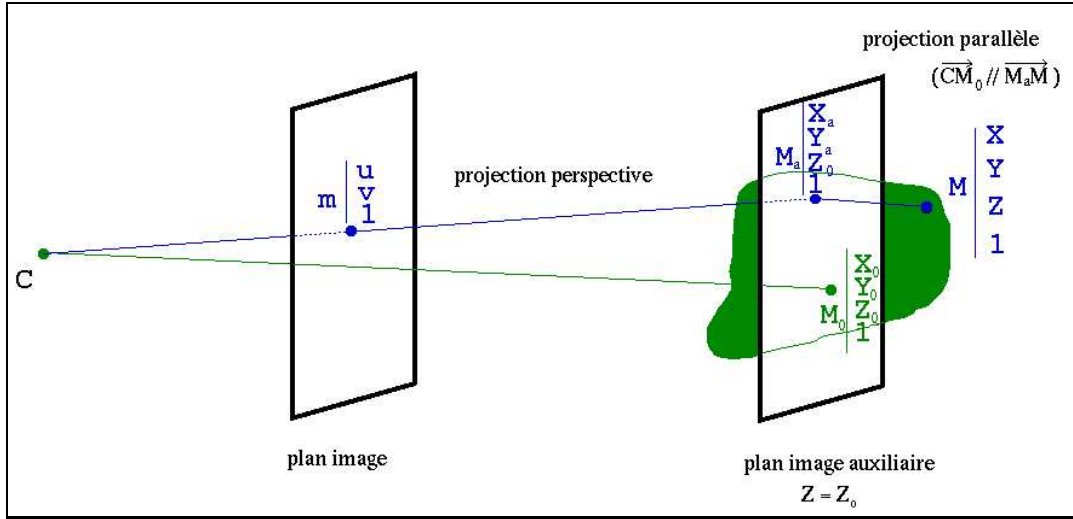


FIG. 3 – Principe de la projection para-perspective

Ecrivons la relation de parallélisme pour la projection parallèle :

$$\begin{cases} \vec{C M_0} \parallel \vec{M_a M} \\ \vec{M_a} \in \mathcal{P}(Z = Z_0) \end{cases} \Leftrightarrow \begin{cases} X_a = X - \frac{X_0}{Z_0} Z + X_0 \\ Y_a = Y - \frac{Y_0}{Z_0} Z + Y_0 \end{cases}$$

Écrivons maintenant les équations de projection perspective :

$$\begin{cases} u = \alpha_u \frac{X_a}{Z_0} + \gamma \frac{Y_a}{Z_0} + u_0 \\ v = \alpha_v \frac{Y_a}{Z_0} + v_0 \end{cases}$$

Soit, en remplaçant X_a et Y_a par leurs expressions :

$$\begin{cases} u = \frac{\alpha_u}{Z_0} X + \frac{\gamma}{Z_0} Y - \frac{1}{Z_0} \frac{\alpha_u X_0 + \gamma Y_0}{Z_0} Z + \alpha_u \frac{X_0}{Z_0} + \gamma \frac{Y_0}{Z_0} + u_0 \\ v = \frac{\alpha_v}{Z_0} Y - \frac{\alpha_v}{Z_0} \frac{Y_0}{Z_0} Z + \alpha_v \frac{Y_0}{Z_0} + v_0 \end{cases}$$

On remarque qu'on obtient alors un modèle linéaire en (X, Y, Z) (la profondeur Z n'apparaît plus au dénominateur) mais, en contrepartie, ce modèle comporte plus de paramètres que le modèle perspectif originel, et ces paramètres dépendent de la scène observée. Sachant qu'il s'agit d'une approximation, on considérera que la variation des paramètres X_0 , Y_0 et Z_0 est faible par rapport aux variations des paramètres internes physiques de la caméra lorsque le mouvement est faible.

En posant :

$$\left\{ \begin{array}{l} \alpha_{u_{para}} = \frac{\alpha_u}{Z_0} \\ \alpha_{v_{para}} = \frac{\alpha_v}{Z_0} \\ \gamma_{para} = \frac{\gamma}{Z_0} \\ \beta_{u_{para}} = -\frac{1}{Z_0} \frac{\alpha_u X_0 + \gamma Y_0}{Z_0} \\ \beta_{v_{para}} = -\frac{1}{Z_0} \frac{\alpha_v Y_0}{Z_0} \\ u_{0_{para}} = u_0 + \alpha_u \frac{X_0}{Z_0} + \gamma \frac{Y_0}{Z_0} \\ v_{0_{para}} = v_0 + \alpha_v \frac{Y_0}{Z_0} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \alpha_u = \alpha_{u_{para}} Z_0 \\ \alpha_v = \alpha_{v_{para}} Z_0 \\ \gamma = \gamma_{para} Z_0 \\ u_0 = u_{0_{para}} - \beta_{u_{para}} Z_0 \\ v_0 = v_{0_{para}} - \beta_{v_{para}} Z_0 \\ \frac{X_0}{Z_0} = \frac{\gamma_{para}}{\alpha_{u_{para}}} \frac{\beta_{v_{para}}}{\alpha_{v_{para}}} - \frac{\beta_{u_{para}}}{\alpha_{u_{para}}} \\ \frac{Y_0}{Z_0} = -\frac{\beta_{v_{para}}}{\alpha_{v_{para}}} \end{array} \right. \quad (2)$$

En sus des paramètres intrinsèques de la caméra interviennent $\frac{X_0}{Z_0}$ et $\frac{Y_0}{Z_0}$, paramètres représentant la direction $g = [X_0 \ Y_0 \ Z_0]^T$ du regard sur la scène, ainsi que Z_0 au titre de facteur d'échelle.

Les équations de la projection para-perspective s'écrivent alors sous la forme :

$$\left\{ \begin{array}{l} u = \alpha_{u_{para}} X + \gamma_{para} Y + \beta_{u_{para}} Z + u_{0_{para}} \\ v = \alpha_{v_{para}} Y + \beta_{v_{para}} Z + v_{0_{para}} \end{array} \right. \quad (3)$$

On remarque que $(u_{0_{para}}, v_{0_{para}})$ représente le projeté dans le plan rétinien du point \mathbf{M}_0 de l'espace, selon une projection perspective et on note également les relations suivantes :

$$\left\{ \begin{array}{l} \beta_{u_{para}} = -\alpha_{u_{para}} X_0 - \gamma_{para} Y_0 \\ \beta_{v_{para}} = -\alpha_{v_{para}} Y_0 \\ u_{0_{para}} = u_0 - \beta_{u_{para}} \\ v_{0_{para}} = v_0 - \beta_{v_{para}} \end{array} \right.$$

On obtient ainsi le modèle para-perspectif qui s'écrit matriciellement de la façon suivante :

$$\mathbf{m} = \begin{pmatrix} \alpha_{u_{para}} & \gamma_{para} & \beta_{u_{para}} & u_{0_{para}} \\ 0 & \alpha_{v_{para}} & \beta_{v_{para}} & v_{0_{para}} \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{M} \quad (4)$$

3.2.2 Approximation au premier ordre

D'un autre point de vue, on peut considérer que l'objet est loin de la scène et écrire la coordonnée de profondeur Z sous la forme :

$$Z = Z_0 (1 + \epsilon_Z + o(\epsilon_Z))$$

ainsi que les autres coordonnées :

$$\begin{aligned} X &= X_0 (1 + \epsilon_X + o(\epsilon_X)) \\ Y &= Y_0 (1 + \epsilon_Y + o(\epsilon_Y)) \end{aligned}$$

En écrivant les équations de projection projective, on obtient :

$$\left\{ \begin{array}{l} u = \alpha_u \frac{X_0}{Z_0} (1 + \epsilon_X + o(\epsilon_X)) \frac{1}{1 + \epsilon_Z + o(\epsilon_Z)} + \gamma \frac{Y_0}{Z_0} (1 + \epsilon_Y + o(\epsilon_Y)) \frac{1}{1 + \epsilon_Z + o(\epsilon_Z)} + u_0 \\ v = \alpha_v \frac{Y_0}{Z_0} (1 + \epsilon_Y + o(\epsilon_Y)) \frac{1}{1 + \epsilon_Z + o(\epsilon_Z)} + v_0 \end{array} \right.$$

Or, il est bien connu que :

$$\frac{1}{1 + \epsilon_Z + o(\epsilon_Z)} = 1 - \epsilon_Z + o(\epsilon_Z)$$

ce qui nous permet de faire disparaître les termes en Z du dénominateur, soit :

$$\begin{cases} u &= \alpha_u \frac{X_0}{Z_0} (1 + \epsilon_X + o(\epsilon_X)) (1 - \epsilon_Z + o(\epsilon_Z)) + \gamma \frac{Y_0}{Z_0} (1 + \epsilon_Y + o(\epsilon_Y)) (1 - \epsilon_Z + o(\epsilon_Z)) + u_0 \\ v &= \alpha_v \frac{Y_0}{Z_0} (1 + \epsilon_Y + o(\epsilon_Y)) (1 - \epsilon_Z + o(\epsilon_Z)) + v_0 \end{cases}$$

Au premier ordre, il vient :

$$\begin{cases} u &= \alpha_u \frac{X_0}{Z_0} (1 + \epsilon_X - \epsilon_Z) + \gamma \frac{Y_0}{Z_0} (1 + \epsilon_Y - \epsilon_Z) + u_0 \\ v &= \alpha_v \frac{Y_0}{Z_0} (1 + \epsilon_Y - \epsilon_Z) + v_0 \end{cases}$$

Or :

$$\begin{cases} \epsilon_X &= \frac{X}{X_0} - 1 \\ \epsilon_Y &= \frac{Y}{Y_0} - 1 \\ \epsilon_Z &= \frac{Z}{Z_0} - 1 \end{cases}$$

D'où les équations suivantes :

$$\begin{cases} u &= \frac{\alpha_u}{Z_0} X + \frac{\gamma}{Z_0} Y - \frac{1}{Z_0} \frac{\alpha_u X_0 + \gamma Y_0}{Z_0} Z + u_0 + \frac{\alpha_u X_0 + \gamma Y_0}{Z_0} \\ v &= \frac{\alpha_v}{Z_0} Y - \frac{1}{Z_0} \frac{\alpha_v Y_0}{Z_0} Z + v_0 + \frac{\alpha_v Y_0}{Z_0} \end{cases}$$

qui sont identiques aux équations (3).

On retrouve ainsi les expressions linéaires en les coordonnées du point \mathbf{M} de l'espace mais dont les coefficients dépendent d'un point "milieu" de la scène \mathbf{M}_0 . On a donc montré l'équivalence entre l'approximation géométrique et l'approximation au premier ordre.

3.3 La projection ortho-perspective

Cette approximation de la projection perspective, inventée par Dementhon et Davis [DD89], consiste en une projection orthographique sur un plan image auxiliaire dont la normale est donnée par l'angle sous lequel le point de profondeur moyenne \mathbf{M}_0 est vu depuis le centre optique de la caméra. La figure 4 illustre ce principe et les notations associées.

Les propriétés de cette projection sont le parallélisme des vecteurs $\mathbf{C} \mathbf{M}_0$ et $\mathbf{M}_a \mathbf{M}$ et l'orthogonalité des vecteurs $\mathbf{M}_a \mathbf{M}_0$ et $\mathbf{M}_a \mathbf{M}$, ce qui se traduit par :

$$\begin{cases} \mathbf{C} \mathbf{M}_0 \parallel \mathbf{M}_a \mathbf{M} \\ \mathbf{M}_a \mathbf{M}_0 \perp \mathbf{M}_a \mathbf{M} \end{cases} \Leftrightarrow \begin{cases} X_0 (X_a - X_0) + Y_0 (Y_a - Y_0) + Z_0 (Z_a - Z_0) = 0 \\ \frac{X - X_a}{X_0} = \frac{Y - Y_a}{Y_0} = \frac{Z - Z_a}{Z_0} \end{cases}$$

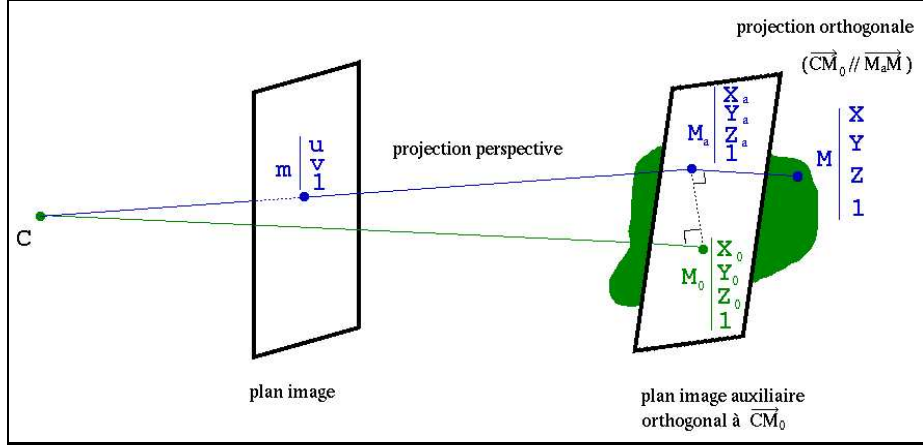


FIG. 4 – Principe de la projection ortho-perspective

Avec un peu de calcul, on en déduit les coordonnées du point auxiliaire \mathbf{M}_a :

$$\begin{cases} X_a = X_0 + X \frac{Y_0 + Z_0}{X_0^2 + Y_0^2 + Z_0^2} - Z \frac{X_0 Z_0}{X_0^2 + Y_0^2 + Z_0^2} - Y \frac{X_0 Y_0}{X_0^2 + Y_0^2 + Z_0^2} \\ Y_a = Y_0 + Y \frac{X_0 + Z_0}{X_0^2 + Y_0^2 + Z_0^2} - X \frac{Y_0 X_0}{X_0^2 + Y_0^2 + Z_0^2} - Z \frac{Y_0 Z_0}{X_0^2 + Y_0^2 + Z_0^2} \\ Z_a = Z_0 + Z \frac{X_0 + Y_0}{X_0^2 + Y_0^2 + Z_0^2} - X \frac{X_0 Z_0}{X_0^2 + Y_0^2 + Z_0^2} - Y \frac{Y_0 Z_0}{X_0^2 + Y_0^2 + Z_0^2} \end{cases}$$

Ainsi, en reportant dans les équations de la projection perspective les expressions des coordonnées de \mathbf{M}_a , on obtient :

$$\begin{cases} u = u_0 + \frac{(-\alpha_u (Y_0^2 + Z_0^2) + \gamma X_0 Y_0) \mathbf{X} + (\alpha_u X_0 Y_0 - \gamma (X_0^2 + Z_0^2)) \mathbf{Y}}{X_0 Z_0 \mathbf{X} + Y_0 Z_0 \mathbf{Y} - (X_0^2 + Y_0^2) \mathbf{Z} - Z_0 (X_0^2 + Y_0^2) - Z_0^3} \\ \quad + \frac{(\alpha_u X_0 Z_0 + \gamma Y_0 Z_0) \mathbf{Z} - (X_0^2 + Y_0^2) (\alpha_u X_0 + \gamma Y_0)}{X_0 Z_0 \mathbf{X} + Y_0 Z_0 \mathbf{Y} - (X_0^2 + Y_0^2) \mathbf{Z} - Z_0 (X_0^2 + Y_0^2) - Z_0^3} \\ v = v_0 + \alpha_v \frac{X_0 Y_0 \mathbf{X} - (X_0^2 + Z_0^2) \mathbf{Y} - (X_0^2 + Y_0^2 + Z_0^2) Y_0}{X_0 Z_0 \mathbf{X} + Y_0 Z_0 \mathbf{Y} - (X_0^2 + Y_0^2) \mathbf{Z} - Z_0 (X_0^2 + Y_0^2) - Z_0^3} \end{cases}$$

Cette approximation, bien qu'intéressante d'un point de vue géométrique n'apporte aucune simplification dans les équations, bien au contraire: elle nécessite des coefficients supplémentaires et le dénominateur des équations est un polynôme du premier degré en les coordonnées tridimensionnelles X , Y et Z .

Au vue de cette complexité et sachant que les conditions de validité de cette projection sont similaires à celles de la projection para-perspective, on préférera cette dernière, à moins, comme dans [Alo90], de considérer une approximation de cette projection ortho-perspective: la projection orthographique qui apporte quant à elle de nombreux avantages et constitue l'objet du paragraphe suivant.

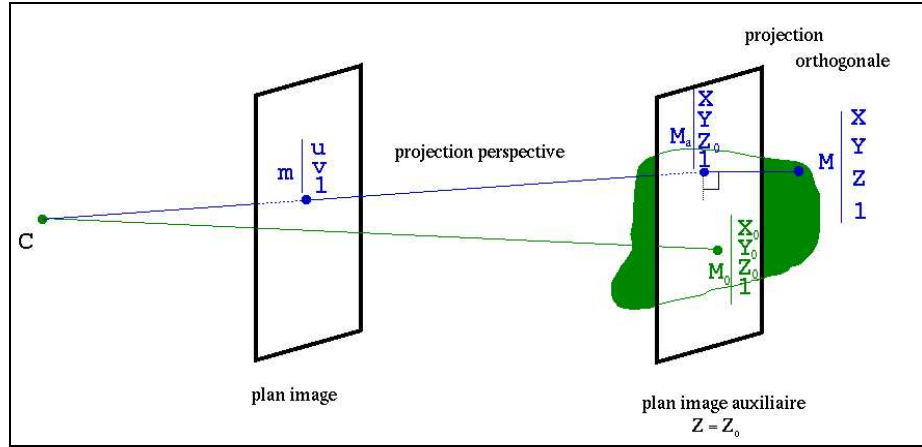


FIG. 5 – Principe de la projection orthographique

3.4 La projection orthographique ou affine

On rencontre également cette approximation sous les noms de perspective faible (*“weak perspective”*) ou orthographique graduée (*“scaled orthographique”*). Elle a été utilisée par de nombreux auteurs comme [Alo90], [CH94], et bien d’autres encore.

Ce modèle est bien adapté à la vision fovéale.

3.4.1 Approximation géométrique

Il s’agit d’un cas particulier à la fois de l’approximation para-perspective et de l’approximation ortho-perspective dans le sens où la projection perspective est approximée par la composée d’une projection orthogonale sur un plan auxiliaire qui est perpendiculaire à l’axe optique; c’est également une projection para-perspective dont la direction est orthogonale au plan image auxiliaire et une projection ortho-perspective dont le plan est orthogonale à l’axe optique de la caméra.

Plus concrètement, il s’agit d’une projection para-perspective pour laquelle la direction du regard et la même que celle de l’axe optique de la caméra.

La figure 5 illustre cette composée et précise les notations.

Les relations géométriques sont les suivantes :

$$\begin{cases} \mathbf{C} \mathbf{M}_0 \parallel \mathbf{M}_a \mathbf{M} \\ \mathbf{M}_a \mathbf{M}_0 \perp \mathbf{M}_a \mathbf{M} \\ \mathbf{M}_a \in \mathcal{P}(Z = Z_0) \end{cases} \Leftrightarrow \begin{cases} X_0 (X_a - X_0) + Y_0 (Y_a - Y_0) + Z_0 (Z_a - Z_0) = 0 \\ \frac{X - X_a}{X_0} = \frac{Y - Y_a}{Y_0} = \frac{Z - Z_a}{Z_0} \\ Z_a = Z_0 \end{cases}$$

Soit :

$$\begin{cases} X_a &= X \\ Y_a &= Y \\ Z_a &= Z_0 \end{cases}$$

Après projection, il vient :

$$\begin{cases} u &= \frac{\alpha_u}{Z_0} X + \frac{\gamma}{Z_0} Y + u_0 \\ v &= \frac{\alpha_v}{Z_0} Y + v_0 \end{cases}$$

On remarque alors qu'on obtient des équations linéaires en les coordonnées X et Y mais, contrairement à l'approximation para-perspective, il n'a pas été nécessaire d'introduire des paramètres supplémentaires. Cependant, les paramètres intrinsèques dépendent, tout comme pour la projection para-perspective, de la profondeur moyenne de la scène.

Ce modèle est équivalent au modèle para-perspectif dans le cas où on regarde droit devant soi (vision fovéale), c'est-à-dire le cas où $X_0 = 0$ et $Y_0 = 0$.

3.4.2 Approximation à l'ordre 0

D'un point de vue équationnel, il s'agit d'un développement à l'ordre 0 en la profondeur Z du modèle perspectif, $Z = Z_0 + o(1)$:

$$\begin{cases} u &= \alpha_u \frac{X}{Z_0} + \gamma \frac{Y}{Z_0} + u_0 \\ v &= \alpha_v \frac{Y}{Z_0} + v_0 \end{cases}$$

Soit, en notant :

$$\begin{cases} \alpha_{u_{ortho}} &= \frac{\alpha_u}{Z_0} \\ \alpha_{v_{ortho}} &= \frac{\alpha_v}{Z_0} \\ \gamma_{ortho} &= \frac{\gamma}{Z_0} \end{cases}$$

sous forme matricielle :

$$\mathbf{m} = \begin{pmatrix} \alpha_{u_{ortho}} & \gamma_{ortho} & 0 & u_0 \\ 0 & \alpha_{v_{ortho}} & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{M} \quad (5)$$

On retrouve bien le modèle orthographique, ce qui nous montre l'équivalence entre l'approximation géométrique et l'approximation à l'ordre 0.

L'avantage de cette projection est de nous procurer, tout comme la projection para-perspective, des équations affines ce qui simplifie énormément les calculs par la suite. Bien plus restreint que le modèle para-perspectif, en général, ce modèle est cependant bien réaliste dans des cas particuliers précis que nous allons étudier dans la suite de ce rapport.

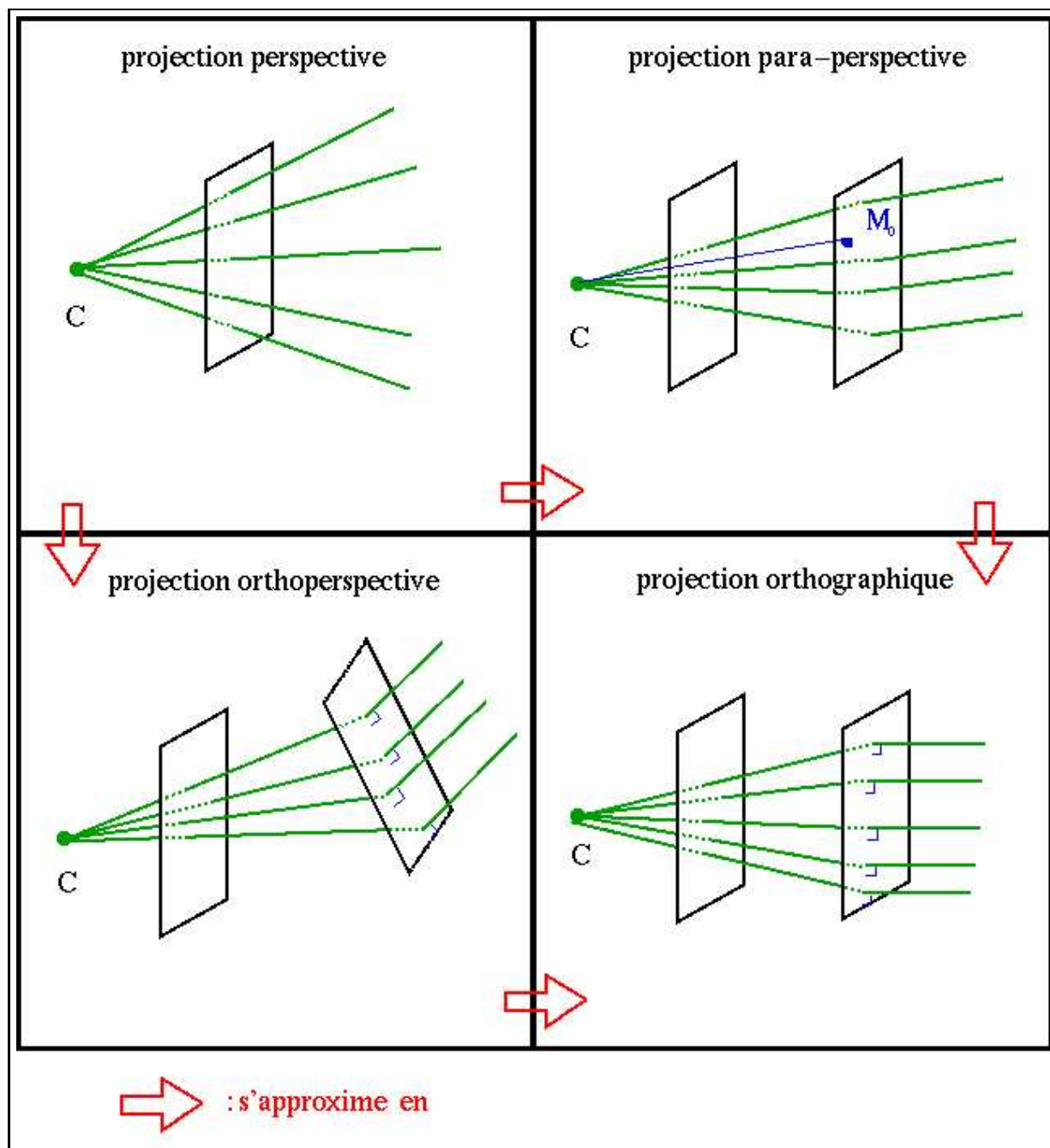


FIG. 6 – La projection perspective et ses approximations.

3.5 Différents modèles en une équation

Nous résumons le principe de la projection perspective et de ses approximations sur la figure 6.

Le modèle perspectif est le plus général. Il peut être approximé, lors de conditions expérimentales similaires en une projection para-perspective ou ortho-perspective. Ces deux approximations peuvent être à leur tour approximées par la projection orthographique.

Compte tenu de la pauvreté en simplification algébrique de la projection perspective, celle-ci n'est utilisée dans la littérature, voir par exemple [Alo90], que dans le cas de son approximation orthographique. D'autre part, des auteurs ayant développés des algorithmes dans le cas spécifique de la projection orthographique, [TK91], [PK93] et [Koe84], ont réussi à généraliser leurs méthodes au cas para-perspectif, ce qui n'était pas faisable pour le cas ortho-perspectif. Enfin, la projection ortho-perspective ayant le même domaine de validité que la projection para-perspective en étant plus compliquée que celle-ci, nous ne considérons, dans la suite que les approximations para-perspective ou orthographique.

Le modèle para-perspectif a l'avantage de linéariser les équations de projection mais possède le désavantage d'introduire deux paramètres supplémentaires. Le modèle orthographique réunit les avantages en linéarisant les équations sans faire intervenir de paramètre supplémentaire.

Ces deux modèles font intervenir la scène par l'intermédiaire d'un point de profondeur moyenne mais conservent tous deux le facteur d'échelle (les objets proches paraissent plus grands que les objets lointains).

Cependant, le modèle para-perspectif reste tout de même plus précis que le modèle orthographique, car il tient compte des positions des objets en périphérie (vus sous un angle différent).

Nous proposons ici un modèle général de projection pouvant englober les modèles vus précédemment (1, 4, 5), en notant \hat{x} la variable générique x :

$$\kappa \mathbf{m} = \begin{pmatrix} \hat{\alpha}_u & \gamma & \lambda \hat{\beta}_u + \mu u_0 & (1 - \mu) \hat{u}_0 \\ 0 & \hat{\alpha}_v & \lambda \hat{\beta}_v + \mu v_0 & (1 - \mu) \hat{v}_0 \\ 0 & 0 & \mu & (1 - \mu) \end{pmatrix} \mathbf{M} \quad (6)$$

Dans ce modèle, on distingue deux types de paramètres :

- **Paramètres modaux** : λ et μ
- **Paramètres intrinsèques** : $\alpha_u, \alpha_v, \beta_u, \beta_v, u_0, v_0$ et γ

Les paramètres modaux $\lambda \in \{0,1\}$ et $\mu \in \{0,1\}$ déterminent le modèle de projection :

	λ	μ
Projection perspective pure	1	1
Projection orthographique	0	0
Projection para-perspective	1	0

Et nous récapitulons dans le tableau suivant les valeurs des paramètres intrinsèques pour les approximations de la projection perspective :

	projection perspective	projection para-perspective	projection orthographique
$\hat{\alpha}_u$	α_u	$\frac{\alpha_u}{Z_0}$	$\frac{\alpha_u}{Z_0}$
$\hat{\alpha}_v$	α_v	$\frac{\alpha_v}{Z_0}$	$\frac{\alpha_v}{Z_0}$
$\hat{\beta}_u$	0	$-\frac{1}{Z_0} \frac{\alpha_u X_0 + \gamma Y_0}{Z_0}$	0
$\hat{\beta}_v$	0	$-\frac{1}{Z_0} \frac{\alpha_v Y_0}{Z_0}$	0
$\hat{\gamma}$	γ	$\frac{\gamma}{Z_0}$	$\frac{\gamma}{Z_0}$
\hat{u}_0	u_0	$u_0 + \alpha_u \frac{X_0}{Z_0} + \gamma \frac{Y_0}{Z_0}$	u_0
\hat{v}_0	v_0	$v_0 + \alpha_v \frac{Y_0}{Z_0}$	v_0

4 Disparité vectorielle entre deux images

On définit la disparité vectorielle entre deux images par :

$$\begin{cases} du &= u_2 - u_1 \\ dv &= v_2 - v_1 \end{cases}$$

représentant la mesure du mouvement rétinien.

Afin de simplifier les notations, on va considérer par la suite que le repère de l'espace et celui de la caméra pour la première image sont confondus. On sait alors se ramener à n'importe quel repère de l'espace par une transformation rigide (rotation et translation), cas particuliers de collinéation. On notera, dans le cas d'un mouvement de la scène, \mathbf{R} et \mathbf{t} la rotation et la translation effectuées par la scène ou, dans le cas d'un mouvement propre de la caméra, \mathbf{R}^T et $-\mathbf{t}$ la rotation et la translation effectuées par la caméra entre les deux images. Si plusieurs objets de la scène sont en mouvement, on ne considèrera dans un premier temps qu'un seul objet à la fois.

La rotation s'exprime comme une exponentielle de la matrice produit vectoriel du vecteur rotation $\mathbf{r} = [r_0 \ r_1 \ r_2]^T$: $\mathbf{R} = (r_{ij})_{i,j} = \exp(\tilde{\mathbf{r}})$ dont la direction et le sens définissent l'axe de rotation et la norme l'angle de rotation. Le vecteur translation s'exprime quant à lui selon : $\mathbf{t} = [t_0 \ t_1 \ t_2]^T$.

4.1 Disparité vectorielle dans le cas de la projection perspective

Dans le cas de la projection perspective, la disparité vectorielle entre deux images s'exprime par :

$$\begin{cases} u &= u_0 + \alpha_u \frac{X}{Z} + \gamma \frac{Y}{Z} \\ v &= v_0 + \alpha_v \frac{Y}{Z} \end{cases}$$

Sachant que les équations du mouvement s'écrivent :

$$\begin{cases} X_2 &= r_{11} X_1 + r_{12} Y_1 + r_{13} Z_1 + t_0 \\ Y_2 &= r_{21} X_1 + r_{22} Y_1 + r_{23} Z_1 + t_1 \end{cases}$$

et en approximant localement une surface par un morceau de plan ne contenant pas l'axe optique :

$$Z_1 = N_X X_1 + N_Y Y_1 + Z_N$$

on obtient une expression de la disparité vectorielle sous la forme :

$$\begin{cases} du &= [C_u + a u + b v + u(e u + g v)]/[C_d - e u - g v] \\ dv &= [C_v + c u + d v + v(e u + g v)]/[C_d - e u - g v] \end{cases}$$

où les valeurs des coefficients peuvent être obtenues par le code Maple suivant :

```
> r:=matrix([[r11,r12,r13],[r21,r22,r23],[r31,r32,r33]]):
> Z := Nx * X + Ny * Y + ZN:
> equ := au*X/Z + gamma*Y/Z + u_0 - u:
> eqv := av*Y/Z + v_0 - v:
> assign(solve({equ,eqv},{X,Y})):
bytes used=23029936, alloc=2096768, time=17.88
> X2:= r[1,1] * X + r[1,2] * Y + r[1,3] * Z + t0:
> Y2:= r[2,1] * X + r[2,2] * Y + r[2,3] * Z + t1:
> Z2:= r[3,1] * X + r[3,2] * Y + r[3,3] * Z + t2:
> u2 := au2 * X2/Z2 + g2 * Y2/Z2 + u0p:
> v2 := av2 * Y2/Z2 + v0p:
> du := factor(u2 - u):
> dv := factor(v2 - v):
> Cu := coeff(coeff(numer(du),u,0),v,0);

Cu := -au2 t0 u_0 Nx av - g2 t1 au v0 Ny + g2 t1 gamma Nx v0 - g2 t1 u_0 Nx av
      - u0p r31 ZN v0 gamma + u0p r31 ZN av u_0 + u0p r32 au ZN v0
      - u0p r33 ZN au av - u0p t2 au v0 Ny + u0p t2 gamma Nx v0 - u0p t2 u_0 Nx av
      - g2 r21 ZN v0 gamma + g2 r21 ZN av u_0 + g2 r22 au ZN v0 - g2 r23 ZN au av
      - g2 t1 au av - u0p t2 au av - au2 r11 ZN v0 gamma + au2 r11 ZN av u_0
      + au2 r12 au ZN v0 - au2 r13 ZN au av - au2 t0 au v0 Ny + au2 t0 gamma Nx v0
      - au2 t0 au av

> Cv := coeff(coeff(numer(dv),u,0),v,0);
Cv := -v0p t2 u_0 Nx av - av2 r21 ZN v0 gamma + av2 r21 ZN av u_0
      + av2 r22 au ZN v0 - av2 r23 ZN au av - av2 t1 au v0 Ny + av2 t1 gamma Nx v0
      - av2 t1 u_0 Nx av - v0p r31 ZN v0 gamma + v0p r31 ZN av u_0
      + v0p r32 au ZN v0 - v0p r33 ZN au av - v0p t2 au v0 Ny + v0p t2 gamma Nx v0
      - av2 t1 au av - v0p t2 au av

> Cd := coeff(coeff(denom(du),u,0),v,0);
Cd := -r31 ZN v0 gamma + r31 ZN av u_0 + r32 au ZN v0 - r33 ZN au av - t2 au av
      - t2 au v0 Ny + t2 gamma Nx v0 - t2 u_0 Nx av

> a := coeff(numer(du),u);
a := g2 t1 Nx av - u0p r31 ZN av + u0p t2 Nx av - r31 ZN gamma v + r31 ZN v0 gamma
      - r31 ZN av u_0 - r32 au ZN v0 + r32 au ZN v + r33 ZN au av + t2 au v0 Ny
      - t2 au v Ny - t2 gamma Nx v0 + t2 gamma Nx v + t2 u_0 Nx av - g2 r21 ZN av
      + t2 au av + au2 t0 Nx av - au2 r11 ZN av

> b := coeff(numer(du),v);
b := g2 r21 ZN gamma + g2 t1 au Ny - g2 t1 gamma Nx + u0p r31 ZN gamma - u0p r32 au ZN
      + u0p t2 au Ny - u0p t2 gamma Nx - u r31 ZN gamma + u r32 au ZN - u t2 au Ny
      + u t2 gamma Nx - g2 r22 au ZN + au2 r11 ZN gamma - au2 r12 au ZN
      + au2 t0 au Ny - au2 t0 gamma Nx

> c := coeff(numer(dv),u);
c := v0p t2 Nx av + v r31 ZN av - v t2 Nx av - av2 r21 ZN av + av2 t1 Nx av
      - v0p r31 ZN av
```

```

> d := coeff(deriv(dv),v);
d := r31 ZN v0 gamma - r31 ZN av u_0 + r31 ZN av u - r32 au ZN v0 + r33 ZN au av
    + t2 au v0 Ny - t2 gamma Nx v0 + t2 u_0 Nx av - t2 u Nx av + av2 r21 ZN gamma
    - av2 r22 au ZN + av2 t1 au Ny - av2 t1 gamma Nx + v0p r31 ZN gamma
    - v0p r32 au ZN + v0p t2 au Ny - v0p t2 gamma Nx + t2 au av

> e := coeff(deriv(du),u^2);
    e := r31 ZN av - t2 Nx av

> g := coeff(deriv(dv),v^2);
    g := r32 au ZN + t2 gamma Nx - t2 au Ny - r31 ZN gamma

```

4.2 Disparité vectorielle dans le cas de la projection para-perspective

On effectue le même raisonnement que précédemment mais en utilisant les équations de la projection para-perspective :

$$\begin{cases} u &= u_0 + \alpha_u X + \gamma Y + \beta_u Z \\ v &= v_0 + \alpha_v Y + \beta_v Z \end{cases}$$

et on obtient alors une expression affine en (u,v) de la disparité vectorielle sous la forme :

$$\begin{cases} du &= C_u + a u + b v \\ dv &= C_v + c u + d v \end{cases}$$

Similairement au paragraphe précédent, on obtient les valeurs des coefficients par quelques calculs :

```

a := coeff(deriv(du),u);
b := coeff(deriv(du),v);
c := coeff(deriv(dv),u);
d := coeff(deriv(dv),v);
> r:=matrix([[r11,r12,r13],[r21,r22,r23],[r31,r32,r33]]);
> Z := Nx * X + Ny * Y + ZN;
> equ := au*X + g*Y + bu*Z + u_0 - u;
> eqv := av*Y + bv*Z + v0 - v;
> assign(solve({equ,eqv},{X,Y})):
bytes used=1000120, alloc=786288, time=0.54
> X2:= r[1,1] * X + r[1,2] * Y + r[1,3] * Z + t0;
> Y2:= r[2,1] * X + r[2,2] * Y + r[2,3] * Z + t1;
> Z2:= r[3,1] * X + r[3,2] * Y + r[3,3] * Z + t2;
>
> u2 := au2 * X2 + g2 * Y2 + bu2 * Z2 + u0p;
> v2 := av2 * Y2 + bv2 * Z2 + v0p;
> du := factor(u2 - u);
> dv := factor(v2 - v);
> Cu := coeff(coeff(deriv(du),u,0),v,0);
Cu := au2 r11 av u_0 - au2 r11 v0 g + au2 r12 au v0 - au2 t0 au av
    + g2 r21 av u_0 - g2 r21 v0 g + g2 r22 au v0 - g2 t1 au av
    + bu2 r31 av u_0 - bu2 r31 v0 g + bu2 r32 au v0 - bu2 t2 au av
    - u0p au bv Ny + u0p g bv Nx - u0p bu Nx av - au2 r11 Ny bu v0
    - au2 r11 bv ZN g + au2 r11 bv Ny u_0 + au2 r11 ZN bu av
    + au2 r12 au bv ZN + au2 r12 bu Nx v0 - au2 r12 u_0 bv Nx
    + au2 r13 Nx av u_0 - au2 r13 Nx v0 g + au2 r13 Ny au v0
    - au2 r13 ZN au av - au2 t0 au bv Ny - u0p au av + au2 t0 g bv Nx

```

```

- au2 t0 bu Nx av - g2 r21 Ny bu v0 - g2 r21 bv ZN g + g2 r21 bv Ny u_0
+ g2 r21 ZN bu av + g2 r22 au bv ZN + g2 r22 bu Nx v0 - g2 r22 u_0 bv Nx
+ g2 r23 Nx av u_0 - g2 r23 Nx v0 g + g2 r23 Ny au v0 - g2 r23 ZN au av
- g2 t1 au bv Ny + g2 t1 g bv Nx - g2 t1 bu Nx av - bu2 r31 Ny bu v0
- bu2 r31 bv ZN g + bu2 r31 bv Ny u_0 + bu2 r31 ZN bu av
+ bu2 r32 au bv ZN + bu2 r32 bu Nx v0 - bu2 r32 u_0 bv Nx
+ bu2 r33 Nx av u_0 - bu2 r33 Nx v0 g + bu2 r33 Ny au v0
- bu2 r33 ZN au av - bu2 t2 au bv Ny + bu2 t2 g bv Nx - bu2 t2 bu Nx av

> Cv := coeff(coeff( numer(dv), u, 0), v, 0);
Cv := -v0p au av + av2 r21 av u_0 - av2 r21 v0 g + av2 r22 au v0
- av2 t1 au av + bv2 r31 av u_0 - bv2 r31 v0 g + bv2 r32 au v0
- bv2 t2 au av - v0p au bv Ny + v0p g bv Nx - v0p bu Nx av
- av2 r21 Ny bu v0 - av2 r21 bv ZN g + av2 r21 bv Ny u_0
+ av2 r21 ZN bu av + av2 r22 au bv ZN + av2 r22 bu Nx v0
- av2 r22 u_0 bv Nx + av2 r23 Nx av u_0 - av2 r23 Nx v0 g
+ av2 r23 Ny au v0 - av2 r23 ZN au av - av2 t1 au bv Ny + av2 t1 g bv Nx
- av2 t1 bu Nx av - bv2 r31 Ny bu v0 - bv2 r31 bv ZN g
+ bv2 r31 bv Ny u_0 + bv2 r31 ZN bu av + bv2 r32 au bv ZN
+ bv2 r32 bu Nx v0 - bv2 r32 u_0 bv Nx + bv2 r33 Nx av u_0
- bv2 r33 Nx v0 g + bv2 r33 Ny au v0 - bv2 r33 ZN au av - bv2 t2 au bv Ny
+ bv2 t2 g bv Nx - bv2 t2 bu Nx av

> a := coeff( numer(du), u);
a := -au2 r11 av - g2 r21 av - bu2 r31 av + au bv Ny - g bv Nx + bu Nx av
- au2 r11 bv Ny + au2 r12 bv Nx - au2 r13 Nx av + au av - g2 r21 bv Ny
+ g2 r22 bv Nx - g2 r23 Nx av - bu2 r31 bv Ny + bu2 r32 bv Nx
- bu2 r33 Nx av

> b := coeff( numer(du), v);
b := au2 r11 g - au2 r12 au + g2 r21 g - g2 r22 au + bu2 r31 g - bu2 r32 au
+ au2 r11 Ny bu - au2 r12 bu Nx + au2 r13 Nx g - au2 r13 Ny au
+ g2 r21 Ny bu - g2 r22 bu Nx + g2 r23 Nx g - g2 r23 Ny au
+ bu2 r31 Ny bu - bu2 r32 bu Nx + bu2 r33 Nx g - bu2 r33 Ny au

> c := coeff( numer(dv), u);
c := -av2 r21 bv Ny - av2 r21 av - bv2 r31 av + av2 r22 bv Nx - av2 r23 Nx av
- bv2 r31 bv Ny + bv2 r32 bv Nx - bv2 r33 Nx av

> d := coeff( numer(dv), v);
d := bu Nx av + au av + au bv Ny - g bv Nx + av2 r21 g - av2 r22 au
+ bv2 r31 g - bv2 r32 au + av2 r21 Ny bu - av2 r22 bu Nx + av2 r23 Nx g
- av2 r23 Ny au + bv2 r31 Ny bu - bv2 r32 bu Nx + bv2 r33 Nx g
- bv2 r33 Ny au

```

4.3 Disparité vectorielle dans le cas de la projection orthographique

De même que précédemment, et en utilisant les équations de la projection orthographique :

$$\begin{cases} u &= u_0 + \alpha_u X + \gamma Y \\ v &= v_0 + \alpha_v Y \end{cases}$$

on obtient encore une expression affine en (u, v) de la disparité vectorielle mais avec des coefficients plus simples et que l'on va pouvoir exploiter :

$$\begin{cases} du &= C_u + a u + b v \\ dv &= C_v + c u + d v \end{cases}$$

avec :

$$\begin{cases} C_u &= u'_0 + \alpha'_u (r_{13} Z_N + t_0) + \gamma' (r_{23} Z_N + t_1) - \frac{u_0}{\alpha_u} (\alpha'_u (r_{11} + N_X r_{13}) + \gamma' (r_{21} + N_X r_{23})) \\ &\quad - \frac{v_0}{\alpha_v} (\alpha'_u (r_{12} + N_Y r_{13}) + \gamma' (r_{22} + N_Y r_{23})) - \frac{\gamma}{\alpha_u} (\alpha'_u (r_{11} + N_X r_{13}) + \gamma' (r_{21} + N_X r_{23})) \\ C_v &= v'_0 + \alpha'_v (r_{23} Z_N + t_1) - \frac{u_0 \alpha'_v}{\alpha_u} (r_{21} + N_X r_{23}) - \frac{v_0 \alpha'_v}{\alpha_v} ((r_{22} + N_Y r_{23}) + \frac{\gamma \alpha'_v}{\alpha_u \alpha_v} (r_{21} + N_X r_{23})) \\ a &= \frac{1}{\alpha_u} [\alpha'_u (r_{11} + N_X r_{13}) + \gamma' (r_{21} + N_X r_{23})] - 1 \\ b &= \frac{1}{\alpha_v} [\alpha'_u (r_{12} + N_Y r_{13}) + \gamma' (r_{22} + N_Y r_{23})] - \frac{\gamma}{\alpha_u} (\alpha'_u (r_{11} + N_X r_{13}) + \gamma' (r_{21} + N_X r_{23})) \\ c &= \frac{\alpha'_v}{\alpha_v} (r_{21} + N_X r_{23}) \\ d &= \frac{\alpha'_v}{\alpha_v} ((r_{22} + N_Y r_{23}) - \frac{\gamma}{\alpha_u} (r_{21} + N_X r_{23})) - 1 \end{cases}$$

Nous avons déjà étudié dans [LV96] le cas où le mouvement est suffisamment faible pour être approximé au premier ordre : $\mathbf{R} = \mathbf{I} + \hat{\mathbf{r}}$. En supposant de plus que le facteur γ est négligeable, et que les paramètres intrinsèques sont fixes, on obtient alors des relations bien plus simples entre les paramètres permettant de retrouver la normale $[N_X \ N_Y \ -1]^T$.

4.4 Retour au cas de la projection perspective

Nous avons vu que dans ce cas, l'expression n'est pas affine :

$$\begin{cases} du &= [C_u + a u + b v + u (e u + g v)] / [C_d - e u - g v] \\ dv &= [C_v + c u + d v + v (e u + g v)] / [C_d - e u - g v] \end{cases}$$

sauf si on annule les coefficients suivants :

$$\begin{cases} e &= t_2 (r_0^2 + r_1^2 + r_2^2 + 4) N_X \alpha_v - 2 Z_N (r_2 r_0 - 2 r_1) \alpha_v \\ g &= -t_2 (r_0^2 + r_1^2 + r_2^2 + 4) \gamma N_X + t_2 (r_0^2 + r_1^2 + r_2^2 + 4) \alpha_u N_Y - 2 \alpha_u Z_N (r_2 r_1 + 2 r_0) \\ &\quad + 2 Z_N (r_2 r_0 - 2 r_1) \gamma \end{cases}$$

Dans le cas d'une configuration générale de points 3D, pour que ces coefficients soient toujours nuls, il faut qu'ils soient indépendants du point (donc de la normale). Ainsi, $t_2 = 0$ et $r_0 = r_1 = 0$ ce qui correspond à un mouvement rétinien (pour lequel le plan rétinien est invariant).

Dans des configurations particulières de points 3D telles que points coplanaires (N_X, N_Y et Z_N constants), ou plus généralement, points tels que $\frac{N_X}{Z_N}$ et $\frac{N_Y}{Z_N}$ soient constants, on peut établir des conditions sur t_2 et r_2 par exemple (à moduler en fonction des annulations de dénominateurs) :

$$\begin{cases} t_2 &= 2 \frac{Z_N}{N_X} \frac{r_2 r_0 - 2 r_1}{r_0^2 + r_1^2 + r_2^2 + 4} \\ r_2 &= 2 \frac{N_Y r_0 - N_X r_1}{N_X r_0 + N_Y r_1} \end{cases} \quad (7)$$

Une application à la robotique du cas de mouvement rétinien est présentée dans [LV96].

On a ainsi montré que dans le cas du mouvement rétinien ou de mouvement défini par l'équation (7) dans des configurations particulières de points, au niveau de la disparité vectorielle, les modèles para-perspectifs ou orthographiques ont des équations similaires à celle de la projection perspective. Les différences entre les deux approximations se situent au niveau des expressions des coefficients.

5 La matrice fondamentale

Après avoir étudié la disparité vectorielle, nous allons maintenant nous intéresser à la matrice fondamentale qui établit une relation entre les points des images sans dépendre de la profondeur comme c'était le cas dans le chapitre précédent.

Nous allons tout d'abord rappeler l'expression de la matrice fondamentale [Fau93] dans le cas de la projection perspective avant de généraliser son expression et sa validité dans les cas para-perspectif et orthographique.

5.1 La matrice fondamentale dans le cas de la projection perspective

Dans ce paragraphe, on notera par $\mathbf{M}_1 = [X_1 \ Y_1 \ Z_1]^T$ un point de l'espace qui se projette en $\mathbf{m}_1 = [u_1 \ v_1 \ 1]^T$ (on prend la partie (3x3) non nulle de la matrice de projection) selon :

$$\kappa_1 \mathbf{m}_1 = \mathbf{A}_1 \mathbf{M}_1 \quad (8)$$

Après une rotation \mathbf{R} et une translation \mathbf{t} , il atteint une position $\mathbf{M}_2 = [X_2 \ Y_2 \ Z_2]^T$ et se projette en $\mathbf{m}_2 = [u_2 \ v_2 \ 1]^T$:

$$\mathbf{M}_2 = \mathbf{R} \mathbf{M}_1 + \mathbf{t} \quad (9)$$

et

$$\kappa_2 \mathbf{m}_2 = \mathbf{A}_2 \mathbf{M}_2 \quad (10)$$

En combinant (9) et (10), on obtient :

$$\kappa_2 \mathbf{m}_2 = \mathbf{A}_2 \mathbf{R} \mathbf{M}_1 + \mathbf{A}_2 \mathbf{t}$$

\mathbf{M}_1 est remplacé par sa valeur en (8) ce qui fournit :

$$\kappa_2 \mathbf{m}_2 = \kappa_1 \underbrace{\mathbf{A}_2 \mathbf{R} \mathbf{A}_1^{-1}}_{\mathbf{H}_\infty} \mathbf{m}_1 + \underbrace{\mathbf{A}_2 \mathbf{t}}_{\mathbf{e}} \quad (11)$$

On remarque l'apparition dans cette équation de l'homographie du plan à l'infini \mathbf{H}_∞ (voir paragraphe 6) ainsi que de l'épipôle \mathbf{e} défini précédemment.

Il suffit alors de prendre le produit vectoriel de $\mathbf{A}_2 \mathbf{t}$ par l'équation (11) et d'en prendre ensuite la produit scalaire par \mathbf{m}_2 pour obtenir :

$$0 = \mathbf{m}_2^T (\mathbf{A}_2 \tilde{\mathbf{t}}) \mathbf{A}_2 \mathbf{R} \mathbf{A}_1^{-1} \mathbf{m}_1$$

La matrice fondamentale, dans le cas de la projection perspective, s'exprime alors par :

$$\mathbf{F} = (\mathbf{A}_2 \tilde{\mathbf{t}}) \mathbf{A}_2 \mathbf{R} \mathbf{A}_1^{-1}$$

ce qui est une expression qui ne dépend pas des coordonnées des points de l'image, contrairement aux coefficients de la disparité vectorielle.

On peut remarquer dans ce cas que la matrice fondamentale ne permettra jamais de retrouver la norme de la translation; celle-ci ne dépend donc que de deux paramètres.

5.2 La matrice fondamentale dans le cas des approximations de la projection perspective

Considérant deux images consécutives (les coordonnées des points sont à nouveau des coordonnées projectives) :

$$\exists(\lambda_1, \lambda_2, \mathbf{M}_1, \mathbf{M}_2, \mathbf{A}_1, \mathbf{A}_2) / \begin{cases} \lambda_1 \mathbf{m}_1 &= \mathbf{A}_1 \mathbf{M}_1 \\ \lambda_2 \mathbf{m}_2 &= \mathbf{A}_2 \mathbf{M}_2 \end{cases} \quad (12)$$

La position \mathbf{M}_2 du point \mathbf{M} est obtenue après une rotation \mathbf{R} et une translation \mathbf{t} depuis la position \mathbf{M}_1 :

$$\mathbf{M}_2 = [\mathbf{R} | \mathbf{t}] \mathbf{M}_1$$

Les équations précédentes s'écrivent :

$$\exists(\lambda_1, \lambda_2, X_1, Y_1, Z_1, \mathbf{A}_1, \mathbf{A}_2) / \begin{cases} \lambda_1 \mathbf{m}_1 &= \mathbf{A}_1 \mathbf{M}_1 \\ \lambda_2 \mathbf{m}_2 &= \mathbf{A}_2 [\mathbf{R} | \mathbf{t}] \mathbf{M}_1 \end{cases}$$

Ce qui peut également se mettre sous la forme :

$$\exists(\mathbf{A}_1, \mathbf{A}_2) / \left| \begin{array}{cc} \mathbf{m}_1 & \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} & \mathbf{A}_1 \\ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} & \mathbf{m}_2 & \mathbf{A}_2 [\mathbf{R} | \mathbf{t}] \end{array} \right| = 0 \quad (13)$$

On peut remarquer qu'il s'agit d'une forme bilinéaire en \mathbf{m}_1 et \mathbf{m}_2 , ce qui montre que ce résultat peut se mettre sous une forme plus habituelle, utilisant une généralisation de la matrice fondamentale :

$$\boxed{\mathbf{m}_2^T \mathbf{F}(\mathbf{A}_1, \mathbf{A}_2, \mathbf{R}, \mathbf{t}) \mathbf{m}_1 = 0} \quad (14)$$

Notons par \mathbf{a}_i les lignes des matrices de projection :

$$\begin{aligned} \mathbf{A}_1 &= [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3]^T \\ \mathbf{A}_2 [\mathbf{R} | \mathbf{t}] &= [\mathbf{a}_4 \ \mathbf{a}_5 \ \mathbf{a}_6]^T \end{aligned}$$

où, pour chaque i : $\mathbf{a}_i = [a_{i1} \ a_{i2} \ a_{i3} \ a_{i4}]$.

Réécrivons l'équation (13) en utilisant cette notation :

$$\left| \begin{array}{ccc} u_1 & 0 & \mathbf{a}_1 \\ v_1 & 0 & \mathbf{a}_2 \\ 1 & 0 & \mathbf{a}_3 \\ 0 & u_2 & \mathbf{a}_4 \\ 0 & v_2 & \mathbf{a}_5 \\ 0 & 1 & \mathbf{a}_6 \end{array} \right| = 0 \quad (15)$$

et notons $d_{\alpha\beta\gamma\zeta}$ le déterminant suivant :

$$d_{\alpha\beta\gamma\zeta} = \begin{vmatrix} \mathbf{a}_\alpha \\ \mathbf{a}_\beta \\ \mathbf{a}_\gamma \\ \mathbf{a}_\delta \end{vmatrix} = \begin{vmatrix} a_{\alpha 1} & a_{\alpha 2} & a_{\alpha 3} & a_{\alpha 4} \\ a_{\beta 1} & a_{\beta 2} & a_{\beta 3} & a_{\beta 4} \\ a_{\gamma 1} & a_{\gamma 2} & a_{\gamma 3} & a_{\gamma 4} \\ a_{\delta 1} & a_{\delta 2} & a_{\delta 3} & a_{\delta 4} \end{vmatrix}$$

En développant le déterminant de l'équation (15) par la première colonne puis par la seconde colonne, on obtient :

$$\mathbf{m}_2^T \underbrace{\begin{pmatrix} d_{2356} & -d_{1356} & d_{1256} \\ -d_{2346} & d_{1346} & -d_{1246} \\ d_{2345} & -d_{1345} & d_{1245} \end{pmatrix}}_{\mathbf{F}} \mathbf{m}_1 = 0$$

Ce qui nous fournit une expression générale de \mathbf{F} , quelque soit la projection utilisée et permet de vérifier, par un simple calcul, que son déterminant est nul. Son expression générale en fonction des paramètres intrinsèques et extrinsèques peut être aisément calculée mais n'apporte aucune clarté.

Cette matrice étant par ailleurs définie à un facteur d'échelle près, comme on peut s'en convaincre par l'équation (14), elle possède ainsi au plus 7 degrés de liberté.

Si on fait le décompte des degrés de liberté :

- \mathbf{R} dépend de 3 paramètres
- \mathbf{t} dépend de 3 paramètres
- \mathbf{A}_1 et \mathbf{A}_2 dépendent généralement de 4 ou 5 paramètres

ce qui fait généralement plus que 7 paramètres: la matrice fondamentale ne permet ainsi pas de retrouver les paramètres dans le cas général mais fournit des relations entre les paramètres appelées équations de Kruppa [Fau93]. Ces équations nous permettent de déterminer certains paramètres si on a une connaissance d'un sous-ensemble de paramètres.

D'autre part, dans le cas des approximations para-perspective et orthographique, on remarque que la dernière ligne de la matrice de projection est $[0 \ 0 \ 0 \ 1]$ ce qui reste vrai après mouvement (\mathbf{R}, \mathbf{t}) donc \mathbf{a}_3 et \mathbf{a}_6 sont identiques et tout déterminant $d_{\alpha\beta\gamma\zeta}$ pour lequel $\{3,6\} \subset \{\alpha,\beta,\gamma,\zeta\}$ est nul. Ainsi, la matrice fondamentale aura pour forme :

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & d_{1256} \\ 0 & 0 & -d_{1246} \\ d_{2345} & -d_{1345} & d_{1245} \end{pmatrix}$$

ce qui n'apporte que peu d'information.

En particulier, dans le cas orthographique, et en supposant γ nul, l'expression de la matrice fondamentale est :

$$\begin{pmatrix} 0 & 0 & -2\alpha_u \alpha_v \alpha'_v (1+t_2)(-r_2 r_1 + 2r_0) \\ 0 & 0 & -2(1+t_2)\alpha_u \alpha_v \alpha'_u (r_2 r_0 + 2r_1) \\ 2\alpha_v \alpha'_u \alpha'_v (r_2 r_1 + 2r_0) & -2\alpha'_u \alpha'_v \alpha_u (-2r_1 + r_2 r_0) & f_{33} \end{pmatrix}$$

où :

$$\begin{aligned} f_{33} = & -2(r_1 r_2 + 2r_0) u_0 \alpha'_u \alpha'_v \alpha_v + v'_0 (2r_1 + r_0 r_2) \alpha_u \alpha'_u \alpha_v \\ & + 2u'_0 (2r_0 - r_1 r_2) \alpha_u \alpha'_v \alpha_v + 2t_0 (2r_0 - r_1 r_2) \alpha_u (\alpha'_u + 1) \alpha'_v \alpha_v \\ & + 2(r_0 r_2 - 2r_1) \alpha_u \alpha'_u \alpha'_v v_0 + 2t_1 (r_0 r_2 + 2r_1) \alpha_u \alpha'_u (\alpha'_v + 1) \alpha_v \end{aligned}$$

On a démontré dans [VL95], dans le cas de la projection perspective, l'équivalence entre une matrice fondamentale de la forme :

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & f_{13} \\ 0 & 0 & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}$$

et un mouvement rétinien (translation parallèle au plan rétinien et rotation d'axe parallèle à l'axe optique). Ceci nous montre donc que les cas para-perspectif et orthographique sont donc des sous-cas du cas de la projection perspective avec mouvement rétinien.

6 Les homographies

6.1 Homographie dans le cas de la projection perspective

Dans le cas où la projection est perspective et la translation \mathbf{t} nulle, la matrice fondamentale est nulle ce qui n'apporte aucune information. Par contre, si on prend les parties (3x3) non nulles de \mathbf{A}_1 et \mathbf{A}_2 , alors, les équations :

$$\begin{cases} Z_1 \mathbf{m}_1 &= \mathbf{A}_1 \mathbf{M}_1 \\ Z_2 \mathbf{m}_2 &= \mathbf{A}_2 \mathbf{M}_2 \\ \mathbf{M}_2 &= \mathbf{R} \mathbf{M}_1 \end{cases}$$

fournissent :

$$\mathbf{m}_2 = \frac{Z_1}{Z_2} \mathbf{A}_2 \mathbf{R} \mathbf{A}_1^{-1} \mathbf{m}_1$$

ce qui constitue une homographie $\kappa \mathbf{m}_1 = \mathbf{H}_\infty \mathbf{m}_2$ entre les deux images $\mathbf{H}_\infty = \mathbf{A}_2 \mathbf{R} \mathbf{A}_1^{-1}$.

D'autres part, si on considère une scène plane où le plan a pour normale \mathbf{n} et distance par rapport à l'origine d , on obtient :

$$[\mathbf{n}^T \mathbf{M}_1 = d] \cap [\mathbf{M}_2 = \mathbf{R} \mathbf{M}_1 + \mathbf{t}] \implies \mathbf{M}_2 = \left(\mathbf{R} + \frac{\mathbf{t} \mathbf{n}^T}{d} \right) \mathbf{M}_1$$

Ce qui correspond au cas de la rotation pure en prenant $\mathbf{R}' = \mathbf{R} + \frac{\mathbf{t} \mathbf{n}^T}{d}$. Ainsi, il y a équivalence entre mouvement sans translation et mouvement d'un plan. Si on se place

dans le cas d'une projection perspective, il existe une relation homographique entre les deux images :

$$\lambda \mathbf{m}_2 = \mathbf{H} \mathbf{m}_1$$

avec :

$$\mathbf{H} = \mathbf{A}_2 \left(\mathbf{R} + \frac{\mathbf{t} \mathbf{n}^T}{d} \right) \mathbf{A}_1^{-1}$$

6.2 Homographie dans le cas de la projection para-perspective

Afin de simplifier les notations, nous noterons les paramètres intrinsèques de la caméra $\alpha_u, \alpha_v, \beta_u, \beta_v, \gamma, u_0$ et v_0 . De plus, pour une matrice \mathbf{A} , on notera par $(\mathbf{A})_3$ son troisième vecteur colonne et par $(\mathbf{A})_{-3}$ la matrice composée des colonnes des \mathbf{A} sauf la troisième colonne. Nous noterons également $\underline{\mathbf{M}} = [X \ Y \ 1]^T$ le vecteur réduit d'un point de l'espace.

Nous avons vu au paragraphe 2 que la projection para-perspective s'exprime de la façon suivante :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} \alpha_u & \gamma & \beta_u & u_0 \\ 0 & \alpha_v & \beta_v & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

ce qui peut s'écrire également :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_{(\mathbf{A})_{-3}} \underbrace{\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}}_{\underline{\mathbf{M}}} + Z \underbrace{\begin{pmatrix} \beta_u \\ \beta_v \\ 1 \end{pmatrix}}_{(\mathbf{A})_3}$$

où $(\mathbf{A})_{-3}$ est une matrice carrée inversible :

$$\det((\mathbf{A})_{-3}) = \alpha_u \alpha_v \neq 0$$

Pour deux vues, les équations de projection s'écrivent alors :

$$\begin{cases} \mathbf{m}_1 &= (\mathbf{A}_1)_{-3} \underline{\mathbf{M}}_1 + Z_1 (\mathbf{A}_1)_3 \Rightarrow \underline{\mathbf{M}}_1 = ((\mathbf{A}_1)_{-3})^{-1} \mathbf{m}_1 - Z_1 ((\mathbf{A}_1)_{-3})^{-1} (\mathbf{A}_1)_3 \\ \mathbf{m}_2 &= \mathbf{A}_2 \mathbf{M}_2 \end{cases} \quad (16)$$

L'équation du mouvement se met sous la forme : $\mathbf{M}_2 = [\mathbf{R}|\mathbf{t}] \mathbf{M}_1$

Ainsi, en combinant ceci avec les équations (16), on obtient :

$$\begin{aligned}
\mathbf{m}_2 &= \mathbf{A}_2 [\mathbf{R}|\mathbf{t}] \mathbf{M}_1 \\
&= (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_{-3} \underline{\mathbf{M}}_1 + Z_1 (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_3 \\
&= \underbrace{(\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_{-3} ((\mathbf{A}_1)_{-3})^{-1}}_{\mathbf{H}_{\infty para}} \mathbf{m}_1 - Z_1 (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_{-3} ((\mathbf{A}_1)_{-3})^{-1} (\mathbf{A}_1)_3 + Z_1 (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_3
\end{aligned}$$

Soit :

$$\mathbf{m}_2 = \mathbf{H}_{\infty para} \mathbf{m}_1 + Z_1 \mathbf{K} \quad (17)$$

où $\mathbf{K} = (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_3 - (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_{-3} ((\mathbf{A}_1)_{-3})^{-1} (\mathbf{A}_1)_3$

L'équation (17) représente une homographie si et seulement si le vecteur $Z_1 \mathbf{K}$ est lui-même issu de \mathbf{m}_1 par une homographie ce qui est équivalent aux deux cas suivants :

1. soit Z_1 est un polynôme du premier degré en X_1 et Y_1 .
2. soit les coordonnées de \mathbf{K} sont nulles,

Nous allons maintenant examiner ces deux cas, le premier correspondant à une contrainte de nature géométrique, le second à une contrainte sur les paramètres de mouvement.

6.2.1 Condition de géométrie :

La première condition est une condition de géométrie des points de la scène et impose que ceux-ci soient sur un plan d'équation $Z = aX + bY + c$, c'est-à-dire un plan ne passant pas par l'axe optique. Le centroïde \mathbf{M}_0 appartenant à ce plan, l'équation sera de la forme :

$$Z = a(X - X_0) + b(Y - Y_0) + Z_0 = \mathbf{n} \underline{\mathbf{M}}$$

Ainsi, l'équation (17) s'écrit :

$$\mathbf{m}_2 = \mathbf{H}_{para} \mathbf{m}_1 + \mathbf{n} \underline{\mathbf{M}}_1$$

et l'équation de projection dans la première image issue de (16) :

$$\begin{aligned}
\mathbf{m}_1 &= (\mathbf{A}_1)_{-3} \underline{\mathbf{M}}_1 + (\mathbf{A}_1)_3 \mathbf{n} \underline{\mathbf{M}}_1 \\
&= \underbrace{((\mathbf{A}_1)_{-3} + (\mathbf{A}_1)_3 \mathbf{n})}_{\mathbf{B}} \underline{\mathbf{M}}_1
\end{aligned}$$

Il nous reste à vérifier que la matrice \mathbf{B} est inversible :

$$\begin{aligned}
\det(\mathbf{B}) &= \begin{vmatrix} \alpha_u + a\beta_u & \gamma + b\beta_u & u_0 + (-aX_0 - bY_0 + Z_0)\beta_u \\ a\beta_v & \alpha_v + b\beta_v & v_0 + (-aX_0 - bY_0 + Z_0)\beta_v \\ 0 & 0 & 1 \end{vmatrix} \\
&= (\alpha_u + a\beta_u)(\alpha_v + b\beta_v) - (\gamma + b\beta_u)a\beta_v \\
&= \alpha_u\alpha_v + a\beta_u\alpha_v + b\beta_v\alpha_u - \gamma a\beta_v
\end{aligned}$$

Dans le cas où le paramètre γ est nul et en utilisant les relations (2), cela revient à imposer :

$$Z_0 \neq a X_0 + b Y_0$$

qui signifie que le plan ne doit pas contenir le centre optique. On veut donc des points éloignés et sur un même plan ne contenant ni l'axe optique ni le centre optique, on veillera donc à ce que ce plan soit un plan approximativement fronto-parallèle.

Dans ce cas, l'homographie s'exprime par :

$$\mathbf{H}_{para} = (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_{-3} ((\mathbf{A}_1)_{-3})^{-1} + \mathbf{n} ((\mathbf{A}_1)_{-3} + (\mathbf{A}_1)_3 \mathbf{n})^{-1}$$

6.2.2 Condition sur le mouvement :

La seconde condition revient à imposer une contrainte sur les paramètres du mouvement. L'expression de l'homographie est simplifiée :

$$\mathbf{H}_{\infty para} = (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_{-3} ((\mathbf{A}_1)_{-3})^{-1}$$

par la nullité du vecteur :

$$\mathbf{K} = (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_3 - (\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_{-3} ((\mathbf{A}_1)_{-3})^{-1} (\mathbf{A}_1)_3 = 0$$

Soit, en développant :

$$\begin{aligned} \mathbf{0} = & \begin{pmatrix} \alpha'_u & \gamma' & \beta'_u & u'_0 \\ 0 & \alpha'_v & \beta'_v & v'_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{13} \\ r_{23} \\ r_{33} \\ 0 \end{pmatrix} \\ & - \begin{pmatrix} \alpha'_u & \gamma' & \beta'_u & u'_0 \\ 0 & \alpha'_v & \beta'_v & v'_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & t_0 \\ r_{21} & r_{22} & t_1 \\ r_{31} & r_{32} & t_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\alpha_u} & \frac{-\gamma}{\alpha_u \alpha_v} & \frac{\gamma v_0 - \alpha_v u_0}{\alpha_u \alpha_v} \\ 0 & \frac{1}{\alpha_v} & \frac{-v_0}{\alpha_u} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_u \\ \beta_v \\ 0 \end{pmatrix} \end{aligned}$$

Dans le produit $\mathbf{A}_2 [\mathbf{R}|\mathbf{t}]$, c'est la dernière colonne qui comprend les paramètres de translation et donc, le vecteur colonne $(\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_3$ ne dépend pas de la translation. Ainsi, dans le produit $(\mathbf{A}_2 [\mathbf{R}|\mathbf{t}])_{-3} ((\mathbf{A}_1)_{-3})^{-1}$, c'est toujours dans la dernière colonne qu'on retrouve les paramètres de translation. Lors du produit de ceci par $(\mathbf{A}_1)_3$ dont la troisième coordonnée est nulle, les termes concernant la translation disparaissent. Ainsi, les contraintes sur le mouvement ne concernent que la rotation. De plus, la dernière coordonnée de \mathbf{K} est nulle. Nous aurons ainsi deux contraintes que l'on peut montrer indépendantes sur les paramètres de rotation: l'axe de rotation est donc fixé.

Afin d'exhiber plus clairement ces contraintes, nous allons tout d'abord considérer que le paramètre γ de la caméra est nul pour les deux images, ce qui constitue une hypothèse tout à fait plausible vue que nous sommes déjà en présence d'une approximation de la projection

perspective. Prenant aussi en compte le fait que $\beta_u = -\alpha_u \frac{X_0}{Z_0}$ et $\beta_v = -\alpha_v \frac{Y_0}{Z_0}$, on obtient alors les contraintes suivantes :

$$\begin{cases} 0 &= -X_0 Z_0 r_0^2 + X_0 Z_0 r_2^2 - 2 X_0^2 r_1 + X_0^2 r_2 r_0 + 2 Y_0 Z_0 r_2 - Y_0 Z_0 r_1 r_0 + 2 Y_0 X_0 r_0 \\ &\quad + Y_0 X_0 r_2 r_1 - 2 Z_0^2 r_1 - Z_0^2 r_2 r_0 \\ 0 &= -2 X_0 Z_0 r_2 - X_0 Z_0 r_1 r_0 - 2 X_0 Y_0 r_1 + X_0 Y_0 r_2 r_0 - Y_0 Z_0 r_1^2 + Y_0 Z_0 r_2^2 + 2 Y_0^2 r_0 \\ &\quad + Y_0^2 r_2 r_1 + 2 Z_0^2 r_0 - Z_0^2 r_2 r_1 \end{cases}$$

Ce qui se met sous la forme :

$$\begin{cases} 0 &= (-r_2 X_0 + r_0 Z_0) ((r_0 Y_0 + 2 Z_0)^2 + (X_0 r_0 + Z_0 r_2)^2 + (Y_0 r_2 - 2 X_0)^2) \\ r_1 &= -\frac{X_0 Z_0 r_0^2 - X_0 Z_0 r_2^2 - 2 Y_0 X_0 r_0 - X_0^2 r_2 r_0 - 2 Y_0 Z_0 r_2 + Z_0^2 r_2 r_0}{2 X_0^2 - Y_0 X_0 r_2 + 2 Z_0^2 + Y_0 Z_0 r_0} \end{cases}$$

La première équation offre deux possibilités :

- soit $r_0 = r_2 \frac{X_0}{Z_0}$ et, par substitution dans la seconde équation, $r_1 = r_2 \frac{Y_0}{Z_0}$
- soit $r_2 = 2 \frac{X_0}{Y_0}$ et $r_0 = -2 \frac{Z_0}{Y_0}$ mais ces valeurs annulent le dénominateur de r_1 . Il n'y a pas de solution dans ce cas.

On obtient donc une solution sous la forme : $\mathbf{r} = \theta \mathbf{M}_0$ L'angle de la rotation est libre mais son axe correspond à la direction du regard.

Un raisonnement similaire peut être fait dans le cas où on prend en compte la non nullité du paramètre γ (voir [Lin99]). L'influence de γ se retrouve sur l'axe de rotation dont la direction est légèrement modifiée.

En résumé, dans le cas de la projection para-perspective, on est en présence d'une relation homographique si et seulement si on respecte une des deux conditions suivantes :

condition sur le mouvement : l'axe de rotation correspond à la direction du regard (direction par laquelle la caméra voit le centroïde de la scène)

condition géométrique : les points de la scène sont sur un plan ne contenant pas l'axe optique

6.3 Homographie dans le cas de la projection orthographique

On emploie les mêmes notations que lors du paragraphe précédent.

Ainsi, nous avons vu précédemment que la projection orthographique s'exprime de la façon suivante :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} \alpha_u & \gamma & 0 & u_0 \\ 0 & \alpha_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

ce qui peut s'écrire également :

$$\mathbf{m} = (\mathbf{A})_{-3} \underline{\mathbf{M}}$$

où $(\mathbf{A})_{-3}$ est une matrice carrée inversible :

$$\det((\mathbf{A})_{-3}) = \alpha_u \alpha_v \neq 0$$

Pour deux vues, les équations de projection s'écrivent alors :

$$\begin{cases} \mathbf{m}_1 &= (\mathbf{A}_1)_{-3} \underline{\mathbf{M}}_1 \Rightarrow \underline{\mathbf{M}}_1 = ((\mathbf{A}_1)_{-3})^{-1} \mathbf{m}_1 \\ \mathbf{m}_2 &= (\mathbf{A}_2)_{-3} \underline{\mathbf{M}}_2 \end{cases} \quad (18)$$

Le mouvement rigide s'exprime quant à lui par :

$$\underline{\mathbf{M}}_2 = ([\mathbf{R}|\mathbf{t}])_{-3} \underline{\mathbf{M}}_1 + Z_1 ([\mathbf{R}|\mathbf{t}])_3$$

En combinant ces équations, on obtient :

$$\mathbf{m}_2 = \underbrace{(\mathbf{A}_2)_{-3} ([\mathbf{R}|\mathbf{t}])_{-3} ((\mathbf{A}_1)_{-3})^{-1}}_{\mathbf{H}_{\infty ortho}} \mathbf{m}_1 + Z_1 \underbrace{(\mathbf{A}_2)_{-3} ([\mathbf{R}|\mathbf{t}])_3}_{\mathbf{K}'}$$

On est donc en présence d'une relation homographique si et seulement si \mathbf{K} est nul ou Z_1 est un polynôme du premier degré en X_1 et Y_1 .

6.3.1 Condition sur le mouvement :

Tout comme dans le cas de la projection para-perspective, on obtient deux conditions indépendantes pour l'existence de l'homographie. On retrouve la condition sur la géométrie de la scène : les points sont situés sur un même plan ne contenant pas l'axe optique. La nullité du vecteur \mathbf{K}' impose des conditions sur les paramètres du mouvement :

$$\mathbf{K}' = \mathbf{0} \Leftrightarrow r_{13} = r_{23} = r_{33} \Leftrightarrow r_0 = r_1 = 0$$

ce qui revient à imposer à l'axe de rotation d'être parallèle à l'axe optique qui, dans le cas de la projection orthographique correspond à la direction du regard. La translation n'est pas contrainte.

On peut alors remarquer que, étant donné les conditions de validité de la projection orthographique vue au paragraphe 3.4, i.e. l'objet est loin de la caméra et le mouvement laisse invariant le plan rétinien, on aura ainsi toujours une relation homographique entre les images.

L'homographie s'exprime par :

$$\mathbf{H}_{\infty ortho} = (\mathbf{A}_2)_{-3} ([\mathbf{R}|\mathbf{t}])_{-3} ((\mathbf{A}_1)_{-3})^{-1}$$

6.3.2 Condition de géométrie :

La seconde condition impose aux points de la scène d'appartenir à un plan ne contenant pas l'axe optique et d'équation :

$$Z = [a \ b \ Z_0] [X \ Y \ 1]^T = \mathbf{n}^T \underline{\mathbf{M}}$$

d'où :

$$\mathbf{m}_2 = ((\mathbf{A}_2)_{-3} ([\mathbf{R}|\mathbf{t}]_{-3} ((\mathbf{A}_1)_{-3})^{-1} + (\mathbf{A}_2)_{-3} ([\mathbf{R}|\mathbf{t}]_3 \mathbf{n}^T ((\mathbf{A}_1)_{-3})^{-1}) \mathbf{m}_1$$

L'homographie s'exprime donc par :

$$\mathbf{H}_{ortho} = ((\mathbf{A}_2)_{-3} ([\mathbf{R}|\mathbf{t}]_{-3} ((\mathbf{A}_1)_{-3})^{-1} + (\mathbf{A}_2)_{-3} ([\mathbf{R}|\mathbf{t}]_3 \mathbf{n}^T ((\mathbf{A}_1)_{-3})^{-1})$$

Il n'est pas étonnant que le cas orthographique s'insère dans le cas para-perspectif car il est un cas particulier de ce dernier pour lequel la direction du regard est celle de l'axe optique.

On peut alors remarquer que, étant donné les conditions de validité de la projection orthographique vue au paragraphe 3.4, i.e. l'objet est loin de la caméra et le mouvement laisse invariant le plan rétinien, on aura ainsi toujours une relation homographique entre les images.

Nous résumons ces résultats dans le tableau suivant :

projection	condition sur le mouvement	condition géométrique
perspective	pas de translation	plan quelconque
para-perspective	axe de rotation \parallel direction du regard	plan* $Z = a(X - X_0) + b(Y - Y_0) + Z_0$
orthographique	axe de rotation \parallel axe optique	plan $Z = aX + bY + Z_0$

*: ce plan ne doit pas contenir le centre optique i.e. $Z_0 \neq aX_0 + bY_0$

7 Conclusion

Cette étude a montré les intérêts des approximations de la projection perspective lorsque celles-ci sont valides. La projection para-perspective, approximation au premier ordre, présente l'avantage de fournir des équations de projection linéaire en les coordonnées des points tridimensionnels mais en ajoutant deux paramètres supplémentaires. La projection orthoperspective est une approximation du même ordre que la para-perspective mais fournit des équations inexploitable ou du moins n'apportant aucune simplification. La projection orthographique, cas particulier à la fois de la projection paraperspective et orthographique, réconcilie ces deux

approches en fournissant des équations linéaires en les coordonnées des points tridimensionnels sans paramètre supplémentaire. On notera toutefois que les paramètres de projection dépendent des coordonnées d'un point central de la scène.

S'intéressant à la disparité vectorielle, on montre que celle-ci est linéaire dans les cas para-perspectif et orthographique, et rationnelle dans le cas perspectif. Cependant, dans le cas perspectif, si le mouvement est rétinien, la disparité est linéaire.

Après avoir généralisé la notion de matrice fondamentale, on s'est intéressé aux mouvements et structures géométriques spécifiques pour lesquels la matrice fondamentale n'est plus utilisable et où il existe une relation homographique entre les points des images.

Ceci est à la base d'un travail bien plus vaste concernant l'étude exhaustive des cas particuliers de mouvement, de modèle de caméra et de structure de scène tridimensionnelle en cours, généralisation de [VL99].

Remerciements

L'auteur souhaite remercier vivement Thierry Viéville pour ses nombreux encouragements.

Table des matières

1	Introduction	3
2	La projection perspective	3
3	Les approximations de la projection perspective	5
3.1	État de l'art	5
3.2	La projection para-perspective	6
3.2.1	Approximation géométrique	6
3.2.2	Approximation au premier ordre	8
3.3	La projection ortho-perspective	9
3.4	La projection orthographique ou affine	11
3.4.1	Approximation géométrique	11
3.4.2	Approximation à l'ordre 0	12
3.5	Différents modèles en une équation	14
4	Disparité vectorielle entre deux images	15
4.1	Disparité vectorielle dans le cas de la projection perspective	15
4.2	Disparité vectorielle dans le cas de la projection para-perspective	17
4.3	Disparité vectorielle dans le cas de la projection orthographique	18
4.4	Retour au cas de la projection perspective	19
5	La matrice fondamentale	20
5.1	La matrice fondamentale dans le cas de la projection perspective	20
5.2	La matrice fondamentale dans le cas des approximations de la projection perspective	21
6	Les homographies	23
6.1	Homographie dans le cas de la projection perspective	23
6.2	Homographie dans le cas de la projection para-perspective	24
6.2.1	Condition de géométrie:	25
6.2.2	Condition sur le mouvement:	26
6.3	Homographie dans le cas de la projection orthographique	27
6.3.1	Condition sur le mouvement:	28
6.3.2	Condition de géométrie:	29
7	Conclusion	29

Références

- [Alo90] J.Y. Aloimonos. Perspective approximations. *Image and Vision Computing*, 8(3):179–192, August 1990.
- [BWW94] Boubakeur Boufama, Daphna Weinshall, and Mike Werman. Shape from motion algorithms: a comparative analysis of scaled orthography and perspective. In J-O. Eklundh, editor, *Proceedings of the 3rd European Conference on Computer Vision*, volume 800 of *Lecture Notes in Computer Science*, pages 199–204, Stockholm, Sweden, May 1994. Springer-Verlag.
- [CH94] Stéphane Christy and Radu Horaud. Euclidian shape and motion from multiple perspective views by affine iterations. Technical Report 2421, INRIA Rhones-Alpes, December 1994.
- [DD89] D. Dementhon and L. S. Davis. Exact and approximate solutions to the three-point perspective problem. Technical Report CAR-TR-471, Computer Vision Laboratory, University of Maryland, 1989.
- [Fau93] O. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.
- [HCD94] R. Horaud, S. Christy, and F. Dornaika. Object pose: The link between weak perspective, para perspective, and full perspective. Technical Report 2356, INRIA, September 1994.
- [HCM97] Radu Horaud, Stéphane Christy, and Roger Mohr. Euclidean reconstruction and affine camera calibration using controlled robot motions. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Grenoble, France, September 1997.
- [HDL97] Radu Horaud, Fadi Dornaika, Bart Lamiroy, and Stéphane Christy. Object pose: The link between weak perspective, paraperspective, and full perspective. *IJCV*, 22(2), 97.
- [Koe84] Koenderinck. The structure of images. *biol. Cybern.*, 50:363–370, 84.
- [Lin99] Diane Lingrand. *Analyse Adaptative du Mouvement dans des Séquences Monoculaires non Calibrées*. PhD thesis, Université de Sophia-Antipolis, INRIA, Sophia, France, 1999.
- [LV95] Diane Lingrand and Thierry Viéville. Dynamic foveal 3D sensing using affine models. Technical Report RR-2687, INRIA Sophia-Antipolis, October 1995.
- [LV96] Diane Lingrand and Thierry Viéville. Dynamic foveal 3D sensing using affine models. In *Proceedings of the International Conference on Pattern Recognition*, volume 1, pages 810–814, Vienna, Austria, August 1996. Computer Society Press.
- [OMS81] Yu-ichi Ohta, Kiyoshi Maenobu, and Toshiyuki Sakai. Obtaining surface orientation from texels under perspective projection. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, pages 746–751, Vancouver, Canada, August 1981.
- [PK93] Conrad J. Poelman and Takeo Kanade. A paraperspective factorization method for shape and motion recovery. Technical Report CMU-CS-93-219, Carnegie Mellon University, School of Computer Science, December 1993.
- [PK94] Conrad J. Poelman and Takeo Kanade. A paraperspective factorization for shape and motion recovery. In J-O. Eklundh, editor, *Proceedings of the 3rd European Conference on Computer Vision*, volume B of *Lecture Notes in Computer Science*, pages 97–108, Stockholm, Sweden, May 1994. Springer-Verlag.
- [Qua96] Long Quan. Self-calibration of an affine camera from multiple views. *IJCV*, 19(1):93–105, May 1996.

-
- [SP95] Stefano Soatto and Pietro Perona. Dynamic rigid motion estimation from weak perspective. In *Proceedings of the 5th International Conference on Computer Vision*, pages 321–328, Boston, MA, June 1995. IEEE Computer Society Press.
 - [TK91] C. Tomasi and T. Kanade. Factoring image sequences into shape and motion. In *Proc. IEEE Workshop on Visual Motion*, pages 21–28, Princeton, NJ, October 1991.
 - [VL95] Thierry Viéville and Diane Lingrand. Using singular displacements for uncalibrated monocular visual systems. Technical Report RR-2678, INRIA Sophia-Antipolis, October 1995.
 - [VL99] Thierry Viéville and Diane Lingrand. Using specific displacements to analyze motion without calibration. *IJCV*, 31(1):5–29, 1999.
 - [ZDFL94] Zhengyou Zhang, Rachid Deriche, Olivier Faugeras, and Quang-Tuan Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence Journal*, 78(1-2):87–119, 1994. Appeared in October 1995, also INRIA Research Report No.2273, May 1994.



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399